Electromagnetism and Special Relativity

by Jolyon Bloomfield

May 14, 2012

1 Overview

Special relativity is a subject that has (in my opinion) often been taught poorly in undergraduate physics, using confusing notation and math, and I think that this more than anything else has contributed to its reputation as a difficult subject. However, at heart, it's really quite a beautiful subject. It has simple but surprisingly profound concepts, and those lead to some rather counterintuitive results. I'll let you in on two secrets. The first is that I can't do special relativity using the math that I was taught when I first learned the subject. The second is that there is a better way! That's why I asked to teach this – because I want to teach you how to use some powerful tools that really bring out the beauty of the theory. Relativity has become the zeitgeist of modern high energy physics, astrophysics, and cosmology, and to me is one of the most amazing pieces of physics that we have. I hope to share that excitement with you, introduce a formalism that really simplifies everything, and show a little of its application.

These notes are the companion notes for a three lecture mini-series on special relativity. Their aim is to complement what is seen in lectures and provide a few more details, mainly because I don't know a good textbook on the subject. However, these notes should not be viewed as a rigorous derivation of all the formulas here, though I will try to point out when I'm being excessively hand-wavy. My primary objective is to provide you with a first immersion in some deeper aspects of special relativity; I don't expect you to memorize all of the equations here. My hope is that the next time you encounter special relativity, these notes will provide a useful reference for you to build upon.

1.1 Prerequisites

These notes assume that you are familiar with the basic concepts of special relativity. I don't assume that you remember any of the math involved with it, and indeed, I intend to show you a beautiful mathematical formalism for simplifying all of the ugly math that your first introduction to special relativity probably entailed. However, I assume that you have come across the following ideas: *Events*, *Length Contraction*, *Time Dilation*, the γ *Factor*, *Lorentz Transformations*, *Proper Time*, *the Addition of Relativistic Velocities*, and *Relativistic Energy and Momentum*. If you haven't seen these ideas before, you will need to read up on them before we start.

1.2 Learning Outcomes

By the end of this mini-series, you should be able to:

- Manipulate index notation
- Lorentz transform scalar, vector, and tensor quantities
- Lorentz transform the Electromagnetic field

1.3 How to use this document

I will cover most of the material in this document in the lectures. Yes, that will make for a fairly rapid pace. As I said above, I don't expect you to understand all of it in detail. What I really want you to understand is what I have described in the *Learning Outcomes* above. Essentially, I want you to really understand two things: component notation and Lorentz transformations. I'm going to show you how electromagnetism works in special relativity, but I don't expect you to be able to derive it. I'll go through the derivation, and at the end, I'll give you some equations. What I want you to be able to do is to use those equations with the presented notation, and perform Lorentz transformations on them. You'll see a lot of formulas regarding electromagnetism in this mini-course, and it's ok if you don't understand them at this point. You're going to spend the rest of this course going over them in fine detail.

I suggest that you read these notes at least once. Hopefully after you've seen the lectures, most of it will make sense. You may like to print off a copy and bring them to lectures to annotate. Scattered throughout these notes are simple exercises for you to do to make sure you understand what is going on – they should only be about 3 lines of math. At the end of this document, there is a problem set for you to really engage with this material.

Good luck, and have fun!

2 Rotations

Fundamentally, special relativity is based on the idea that what we see as space and time may be different to what somebody else sees, depending on how fast we are traveling relative to each other. It turns out that to transform between reference frames, all you need to do is *rotate*, but in a special way. We'll start with rotations in three dimensions, before expanding to consider an abstract rotation in four dimensions.

I'm going to start in three-dimensional Euclidean space. That's normal three-dimensional space, with no time dimension. Most of this should be familiar to you, but the notation I'm going to use will probably be new. Consider a vector in this space, like $\mathbf{r}$. If we were to write it out completely, we'd say $\mathbf{r} = x\hat{x} + y\hat{y} + z\hat{z}$. Or, you could write it like

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{1}$$

as a column vector. It could be a position vector, a velocity vector, a force vector, it doesn't matter – they're all just vectors in our space.

I'd like to consider rotating this vector. There are two ways to think about this. The first is that we actually change the vector. That is, we pick it up and swing it around so that it points in a different direction. That's probably what you think of when you think of a rotation. However, there is a second way to think about this manipulation. We could also be rotating *ourselves* (as well as our coordinate system), and needing to update our coordinates for the vector because our axes have been rotated. This is the way that I want you to think of rotations in this context – changing reference frames.

Now, one of the things you may have seen in linear algebra is that you can rotate a column vector by applying a matrix to it. As I discussed above though, I'm not going to rotate my column vector, what I want to rotate is the coordinate system that I'm using. I'm going to turn myself around an axis by some angle, and then ask "If I set up my new coordinate system in exactly the same way relative to myself as before I rotated, what are the coordinates on that vector?" Here's an example of a rotation around the z-axis:

$$\mathbf{r}' = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2}$$

Here, we rotate by θ about the z-axis in a counterclockwise direction (assuming we're looking "down" on the xy plane), and end up with a new vector, $\mathbf{r}'$, which describes the vector in my new coordinate system.

Vectors, like $\mathbf{r}$, need to transform if we change our coordinate system. You should be fairly happy with these statements so far. There are three axes that we can rotate around, and so for you mathematicians out there, this forms the $SO(3)$ rotation group.

Now consider the quantity $\mathbf{r} \cdot \mathbf{r}$. How does this transform? I leave it as an exercise for you to show that if we transform $\mathbf{r}$ to $\mathbf{r}'$, we will find that

$$\mathbf{r} \cdot \mathbf{r} = \mathbf{r}' \cdot \mathbf{r}'. \tag{3}$$

This happens because $\mathbf{r} \cdot \mathbf{r}$ is a *scalar* quantity. You probably know that a scalar is a number, unlike a vector, which has magnitude and direction. Consider the following question: is a component of a vector a scalar? It's just a number, after all. The answer is, surprisingly enough, no. There is more to being a scalar than just being a number! The extra property that describes a scalar is that it doesn't transform under a rotation. If you think about the rotation of our quantity $\mathbf{r} \cdot \mathbf{r}$ physically, that's kind of obvious – the length of your vector doesn't change if you rotate it, it just points to a different position on the same sphere, centered on the origin. However, the x component of the vector $\mathbf{r}$ does change when you rotate it, and so it is *not* a scalar quantity. We'll discuss this property of scalars in some detail later on.

I'm going to introduce some new notation here. Instead of referring to the entire vector $\mathbf{r}$, it turns out to be really useful to refer to its components, r^i. Here, i could be x, y, or z. To calculate $\mathbf{r} \cdot \mathbf{r}$, we'll write it as $\sum_{i,j} r^i r^j \delta_{ij}$, where δ_{ij} is the Kronecker delta, which is 1 if $i = j$, and 0 if $i \neq j$. Let's write that out completely.

$$\mathbf{r} \cdot \mathbf{r} = r^x r^x + r^y r^y + r^z r^z \tag{4}$$

$$= \sum_i r^i r^i \tag{5}$$

$$= \sum_{i,j} r^i r^j \delta_{ij} \tag{6}$$

$$= r^i r^j \delta_{ij} \tag{7}$$

You may be wondering how I got to that last line. In relativity, summations over all the components of an expression are so common, that it quickly becomes tedious to write out all the summation signs. So, in the last line, I have introduced a convention called the *Einstein Summation Convention*, where repeated indices (i and j here) are summed over. Any index that appears once as a raised index and once as a lowered index has an implicit summation out the front, which runs over all of the coordinates.

We call δ_{ij} the *metric* on Euclidean space. It tells us how to take the dot product of vectors in a coordinate system. (It also does a whole lot of other things, but we'll get to them later.)

Now, note that when we did our rotation above, we used a matrix. We can write a matrix in component form, too. Note that we wrote our vectors (column vectors) with raised indices. It turns out to be really convenient if we write a row vector as a vector with lowered indices. Then, a matrix is going to have one raised and one lowered index. Let's write the rotation matrix above as $R^{i'}{}_i$, where the raised index tells us which row we're on, and the lowered index tells us which column we're on. Then, we can write our transformed vector as the following:

$$r^{i'} = R^{i'}{}_i r^i \tag{8}$$

where again, we've used the Einstein summation convention (summing over i).

Ok, let's keep on going with this notation, and see what we can find out about rotations. Firstly, we should figure out how to rotate a row vector. Say we have a row vector, r_i. How do we rotate it? Well, let's consider the quantity $(\underline{R}\mathbf{r})^T$. From linear algebra, we know that this is equal to

$$(\underline{R}\mathbf{r})^T = \mathbf{r}^T \underline{R}^T \tag{9}$$

The column vector becomes a row vector, and our rotation matrix becomes transposed. However, because our rotation matrix is an orthogonal matrix, its transpose is equal to its inverse, and so we have

$$(\underline{R}\mathbf{r})^T = \mathbf{r}^T \underline{R}^{-1}. \tag{10}$$

Thus, we find that we need to use the inverse rotation matrix to transform row vectors!

$$r_{i'} = R^i{}_{i'} r_i \tag{11}$$

Take note of the notation I'm using here. $R^{i'}{}_i$ is a rotation matrix, and $R^i{}_{i'}$ is its inverse. I'll typically use primes to denote coordinates in a different reference frame, and so by their position in our matrix, you can tell if we're using the rotation matrix or its inverse.

What happens when a matrix multiplies its inverse? We have to get the identity matrix, because we're just unrotating the rotation.

$$R^i{}_{i'} R^{i'}{}_j = \delta^i_j \tag{12}$$

Here, δ^i_j is a representation of the identity matrix. Again, Einstein Summation Convention is employed.

Ok, let's look at the metric now. The metric tells us how to take dot products of vectors. However, it only knows how to take dot products of vectors in our original frame. We'll need to tell it what the x coordinate looks like in the new coordinate system! Well, the metric has two lowered indices, so we'll guess that both indices need to be rotated. Indeed, that is how it is done, and we can write

$$\delta_{i'j'} = \delta_{ij} R^i{}_{i'} R^j{}_{j'}. \tag{13}$$

(For those of you who remember your linear algebra, if we write the metric as a matrix, we're doing a unitary transformation on the matrix.) It turns out that the new metric is exactly the same as the old metric, which is kind of special, but exactly what we would expect from rotating our reference frame – the coordinate system is still Cartesian, so it should be exactly the same.

Now, let's see something special. Remember $\mathbf{r} \cdot \mathbf{r}$? Let's show that it doesn't change if you perform a rotation. Here, we need to rotate all three quantities properly (two vectors and the metric).

$$\begin{aligned}
\mathbf{r}' \cdot \mathbf{r}' &= r^{i'} r^{j'} \delta_{i'j'} && (14)\\
&= R^{i'}{}_i r^i R^{j'}{}_j r^j \delta_{kl} R^k{}_{i'} R^l{}_{j'} && (15)\\
&= \left(R^k{}_{i'} R^{i'}{}_i\right)\left(R^l{}_{j'} R^{j'}{}_j\right) r^i r^j \delta_{kl} && (16)\\
&= \delta^k_i \delta^l_j r^i r^j \delta_{kl} && (17)\\
&= r^i r^j \delta_{ij} && (18)\\
&= \mathbf{r} \cdot \mathbf{r} && (19)
\end{aligned}$$

Note on the second line, we needed more letters for all of the indices, so I had to switch to using k and l. That's ok; they're only dummy indices anyway, just like in integration (you can call them whatever you want). Because we're using this component notation, we don't need to worry about matrix multiplication rules – all of the indices line up just the way they should. On the fourth line, we get some identity matrices (confusingly, also called Kronecker deltas). These identity matrices are really easy to use – they just substitute in one index for another, like $\delta^i_j r^j = r^i$, and above, we used them on the metric. Using this component notation, we see that the length of a vector doesn't change under a rotation.

So far, we've considered ordinary rotations ("rotations on the sphere"), keeping $\mathbf{r}^2 = x^2 + y^2 + z^2$ constant. What we would like to consider is "rotations on the hyperboloid", keeping $s^2 = -t^2 + x^2 + y^2 + z^2$ constant, where I've suggestively called our new coordinate t. What is different between s^2 and $\mathbf{r}^2$? We can write $s^2 = -t^2 + \mathbf{r}^2$, so we've got a four-dimensional space, and we've got a funny negative sign on the t component when we square s. If we write s^μ as a vector, we'll have $s^\mu = (t, x, y, z)$, or $s^\mu = (t, \mathbf{r})$ for short. Keep in mind that s^μ will be a column vector, but I've just written it like this because it's quicker and easier, and you know it's a column vector because of the raised index. I'm using Greek indices for four dimensions instead of Roman indices, which I will only use for spatial indices. We'll let 0 be the t component in our component notation.

When we calculate s^2, we're going to need a four-dimensional metric to take the dot product, and it won't just be a four-dimensional Kronecker delta, because of that minus sign. We can write as before,

$$s^2 = g_{\mu\nu} s^\mu s^\nu \tag{20}$$

where our metric here, which we call $g_{\mu\nu}$, is a diagonal metric, with entries $(-1, 1, 1, 1)$ on the diagonals. This is called the Minkowski metric. (Note that a lot of books will use $(1, -1, -1 - 1)$ instead; this is just a choice of convention. Don't be confused!)

Now, let's think of the transformations that will keep s^2 invariant. Obviously, anything that will keep $\mathbf{r}^2$ invariant and doesn't change the t component will also leave s^2 invariant. Here's an example. I call it Λ instead of R, because it's in four dimensions, and we should keep them separate. This is a normal rotation matrix on the appropriate space coordinates, just like before.

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{21}$$

Like the previous matrix, this just represents a rotation around the z axis. We can do similar transformations for the x and y axes. If we want to transform our vector (now a four-dimensional vector, or *four-vector*), we write $s^{\mu'} = \Lambda^{\mu'}{}_{\mu} s^{\mu}$. Try it – you'll see that the rotation is performed, just as expected.

Ok, but how do we rotate in the t dimension when the metric has a minus sign? We can consider a simpler problem in just two coordinates, t and x. What transformation leaves $-t^2 + x^2$ invariant? The curve $-t^2 + x^2 = C$ describes a hyperbola. To rotate from one point on the hyperbola to another, we can use a hyperbolic transformation.

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta \\ -\sinh\eta & \cosh\eta \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \tag{22}$$

As an exercise, check to see that $-t^2 + x^2 = -t'^2 + x'^2$.

Back in four dimensions, we do exactly the same thing, but generalize it a little bit. Here's exactly the same rotation matrix, just in four dimensions.

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{23}$$

This is a rotation in the tx plane in Minkowski space. You should be able to guess what the rotations in the ty and tz planes look like. We now have three normal rotations, and three special rotations into time. For you mathematicians, this enhanced rotation group is called $SO(3, 1)$ because of that minus sign in the metric (if it was four $+1$s, then the rotation group would be $SO(4)$). So, we now know what the six rotations look like in Minkowski space.

Now that we've developed this machinery, what can we do with it? Let's see some physics!

3 Lorentz Transformations

The amazing insight of Einstein was to realize that space and time were linked. You should have seen previously that they kind of merge together when you perform a Lorentz transformation. What we would like to do is to describe all of this using our machinery from above.

You should be familiar with the concept of an *event*, a location in space and time (often described in some particular reference frame). We'll call it x^μ, where μ can be either t, x, y, or z. Often, we'll let $t = 0$, $x = 1$, $y = 2$, and $z = 3$, just to make summation easier. If we were to write out all of the components of an event, it would look something like the following: $x^\mu = (t, x, y, z)$. Note that this is a vector – we're giving it a raised index, and we would write it as a column if we wrote it out in full. Now, here's a slight difficulty. The components of x^μ have different units! We really can't have that, so we multiply the t by the speed of light, c, to restore the units to what we want them to be. The speed of light is, after all, the same in all reference frames by one of the postulates of special relativity[1]. However, it gets really annoying to bring that c along, so we're going to set $c = 1$. If you ever want to restore c, just use dimensional analysis. If this seems strange, you may want to look up the Buckingham Pi theorem, which justifies doing this.

You may have come across the concept of the Lorentz invariant $s^2 = -t^2 + x^2 + y^2 + z^2$ in a previous special relativity course. If not, that's ok. The idea of this invariant is that it's the same thing in all reference frames, and we saw the rotations that preserve it in the previous section. It turns out that those rotations in Minkowski space from the previous section are just Lorentz transformations in disguise!

Let's see how this works. Recall this Lorentz transformation that we had above.

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{24}$$

If we want to translate our event to a different reference frame, we'd write $x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^{\mu}$. To figure out what the transformation is doing, let's look at a particular event, $x^\mu = (t, 0, 0, 0)$. In the new reference frame, you should find that $x^{\mu\prime} = (t\cosh\eta, -t\sinh\eta, 0, 0)$. This describes a reference frame that is traveling along the x axis at a velocity greater than zero (for positive η) compared to our original frame. We can tell this because what we see as the spatial origin, they see as a point receding into the negative x direction. This is an important point, and we did some physics thinking to find it. Make sure you can work this out by yourself. We call this transformation a *boost* in the x direction.

The parameter η is called the *rapidity* of the transformation. We can restore this to looking like a normal Lorentz transformation in terms of a velocity v by letting

$$\gamma = \cosh\eta \tag{25}$$

$$v\gamma = \sinh\eta \tag{26}$$

$$v = \tanh\eta \tag{27}$$

where v is the boost velocity in the x direction, and $\gamma = 1/\sqrt{1 - v^2}$. This would give us

$$\Lambda^{\mu'}{}_{\mu} = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{28}$$

as you should recognize. (If not, try acting this Lorentz transformation on the event vector x^μ.) Note that inserting a negative η gives you a boost in the $-x$ direction.

So, we now have an idea about how to transform events into different reference frames – we use Lorentz transformations, which are just rotations in disguise. Now, we want to understand how our Lorentz transformations work a little better.

In direct analogy to our three-dimensional case above, we find that we have to transform row vectors using the inverse Lorentz transformation:

$$x_{\mu'} = \Lambda^{\mu}{}_{\mu'} x_{\mu}. \tag{29}$$

The Lorentz transformation, multiplied by the inverse Lorentz transformation, has to give us back the identity matrix by definition, and so we have

$$\Lambda^{\mu}{}_{\mu'} \Lambda^{\mu'}{}_{\nu} = \delta^{\mu}_{\nu}. \tag{30}$$

This makes sense, because if we boost forwards, and then boost backwards, we should end up with what we started with. Finally, if we act on the metric with a Lorentz transformation, we need to do the same thing that we did for the metric in Euclidean space – transform both of the lowered indices.

$$g_{\mu'\nu'} = \Lambda^{\mu}{}_{\mu'} \Lambda^{\nu}{}_{\nu'} g_{\mu\nu} \tag{31}$$

[1]Einstein originally postulated special relativity in this way. However, you can actually show that it is sufficient to require that physics behaves the same way in all directions, at all points in spacetime, for all observers in inertial reference frames. This extra "postulate" will come out of these requirements.

Again, as you will find if you calculate it out, the metric doesn't change under Lorentz transformations. This is one of the defining postulates of special relativity: that the equations of physics are the same in all directions, in all reference frames. Thus, we expect that the metric doesn't change under a Lorentz transformation.

Ok, we've stated that s^2 is invariant under a Lorentz transformation, but I'd like you to derive it yourself, using our component notation, and the properties of the Lorentz transformations that we've described. You can follow exactly the same work as we did for the rotation of $\mathbf{r}$ above. What you want to show is that

$$s^2 = x^\mu x^\nu g_{\mu\nu} = x^{\mu'} x^{\nu'} g_{\mu'\nu'} \tag{32}$$

for any two reference frames. Good luck!

4 Velocity

We should be able to get a velocity vector by differentiating x^μ with respect to time, so let's give that a try.

We need to be really careful when we differentiate with respect to time. What we want to do is to say that t, x, y and z are functions of some parameter (called an *affine parameter*, in case you were curious), and differentiate with respect to that parameter. However, we want that parameter to be time, except that not everybody agrees on the same definition of the time coordinate t. The parameter that makes the most sense to use is called *proper time*. Proper time is the time that an observer in their own frame measures. If I'm looking at you in a different rest frame, then your time will be passing at a different rate compared to my time. This is the famous *time dilation*. The relationship (which I'm not going to derive, but you can do so using Lorentz transformations if you so desire) is the following:

$$t = \tau\gamma. \tag{33}$$

Here, t is my coordinate time. τ (always used for proper time) is your proper time, and γ is the gamma function for the velocity at which you're traveling. Now that we have that, we can calculate the *four-velocity*.

$$v^\mu = \frac{dx^\mu}{d\tau} \tag{34}$$
$$= \left(\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}\right) \tag{35}$$
$$= \frac{dt}{d\tau}\left(1, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \tag{36}$$
$$= \gamma\,(1, \mathbf{v}) \tag{37}$$

In the third line, we took out a factor of $dt/d\tau$, and used the chain rule for the x, y and z components. In the fourth line, we used our above expression to find $dt/d\tau$, and we also introduced the three-velocity, $\mathbf{v}$. Note that this makes sense; we want the three-velocity to somehow be involved in how fast we observe the event to be moving.

Having defined our velocity now, we can calculate: what is the magnitude of the four-velocity?

$$v^2 = v^\mu v^\nu g_{\mu\nu} \tag{38}$$
$$= (-1 + \mathbf{v}^2)\gamma^2 \tag{39}$$
$$= -1 \tag{40}$$
$$= -c^2 \tag{41}$$

In the last line, I reintroduce the factor of the speed of light just for clarity. We see a few rather unusual things here. The first is that because of our choice of metric (called the *signature* of the metric), a quantity squared can be negative. That's ok, because we're not in Euclidean space any more, Dorothy. The second unusual thing we see is that v^2 is a constant, $-c^2$! This implies to us that whatever our four-velocity is, we're "moving" through spacetime at the speed of light!

What happens to our velocity four-vector when we transform to another reference frame? Just as in three dimensions, our rotation matrices could be used to rotate any vector, in four dimensions, Lorentz transformations can be used to transform any four-vector.

$$v^{\mu'} = \Lambda^{\mu'}{}_{\mu} v^{\mu} \tag{42}$$

Be careful now, there are two velocities in this statement. One is the velocity of our original vector, and the other is the velocity of the new frame. If you do the calculation, you'll find the relativistic addition of velocity formula that you should be familiar with. If you're like me, the derivation for this using old notation was somewhat akin to black magic. The fact that we get relativistic addition of velocity just out of our notation should help you appreciate the power of this formalism[2].

There's one more thing to note that's important about v^2. Note that it doesn't depend on what frame we're in. That's because it's a scalar, which doesn't change under Lorentz transformations, just like $\mathbf{r}^2$ in Euclidean space.

5 Tensors

The word "tensor" has a stigma attached to it, the idea that it's a difficult concept. It's a bit more general than a vector or a matrix, but this doesn't mean it's really any harder. In this section, I want to tell you what a tensor is, and we'll look at how to Lorentz transform it.

A tensor is a mathematical object that exists without a coordinate system. However, if we want to work with it, we need to put it into a coordinate system to use it. The tensor is the same object, regardless of what coordinate system we put it in, but our representation of the tensor will change between coordinate systems.

See, that wasn't so bad, was it? Let's look at some different tensors.

The simplest form of tensor is a *scalar*. A scalar is special, because it takes on the same value *regardless of what coordinate system it is in.* Here's some examples of a scalar: the number of apples in a box, the number of electrons in a rock, the total electric charge of the earth. A scalar is just a number that doesn't change under coordinate transformations. You've come across some examples already. The quantity v^2 is a scalar. The invariant s^2 is a scalar. They don't transform. They're the same, however you look at them.

Let's go to the next simplest tensor, a vector. Here's an example: x^μ, the position four-vector. Can you think of another? What about v^μ, the velocity four-vector? When we transform vectors, we need to use a Lorentz transformation. Exercise: write down the Lorentz transformation for each of these vectors. These tensors need to transform – they change when we use a different coordinate system. We can also have the equivalent of a row vector, x_μ. This is a tensor too, and we've told you above how it transforms. Write that down, too.

More complicated tensors have more indices, called Lorentz indices. You've already met one, the metric $g_{\mu\nu}$. Remember how we could write that as a 4×4 matrix? We could have really complicated tensors, like $T^{\mu\nu}{}_{\lambda\sigma}{}^{\rho}$. This one is too ugly to write as a matrix – it would need to be $4 \times 4 \times 4 \times 4 \times 4$! If we have to work with that many indices, we just write out all of the components individually. The important thing is that they all transform in exactly the same way under a Lorentz transformation. You need one factor of Λ for each index, and you need to make sure that you have the correct transformations and inverse transformations as necessary. So, our example here would transform as

$$T^{\mu'\nu'}{}_{\lambda'\sigma'}{}^{\rho'} = T^{\mu\nu}{}_{\lambda\sigma}{}^{\rho} \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} \Lambda^{\lambda}{}_{\lambda'} \Lambda^{\sigma}{}_{\sigma'} \Lambda^{\rho'}{}_{\rho}. \tag{43}$$

There are some special tensors that do not change when you transform them. You've met two already: the metric, $g_{\mu\nu}$, and the identity matrix[3], δ^μ_ν.

[2]Hang on, you might ask, where's the formula? Well, you can calculate it if you want to. However, I actually don't so much care what the formula is; I care that I understand where it comes from. You will derive this formula in the problem set at the end of these notes.

[3]It turns out that there are only three tensors that don't change under Lorentz transformations. We won't need the third one here, but I include some details of it in the appendix just to keep you happy.

You've already been using tensors without knowing it, and you already know how to transform tensors. So don't let the name intimidate you!

Now that we know about tensors, what can we do with them?

6 Using Tensors

There are a few important things to know about using tensors. The first thing we'd like to do is to relate column and row vectors. We can form a row vector from a column vector by using the metric:

$$x_\mu = g_{\mu\nu} x^\nu. \tag{44}$$

We can similarly go the other way, this time using the inverse metric, which as a matrix, is exactly the same as the metric: it's diagonal, with entries $-1, 1, 1, 1$ on the diagonals. This looks like:

$$x^\mu = g^{\mu\nu} x_\nu. \tag{45}$$

Actually, we can use the metric to raise or lower any index. For example,

$$f_{\mu\nu} = g_{\mu\lambda} f^\lambda{}_\nu. \tag{46}$$

Given that we can use the metric to raise and lower indices like this, it becomes really important to make sure that we keep the right indices in the right locations. The metric tensor $g_{\mu\nu}$ is symmetric in μ and ν, but other tensors might not be. What happens when we raise an index on the metric? Go ahead, compute it. You should find that

$$g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_\lambda. \tag{47}$$

One of the special things we can do with raising and lowering indices is swap them, as in the following example.

$$x^\mu v_\mu = x^\mu v^\nu g_{\mu\nu} = x_\nu v^\nu = x_\mu v^\mu \tag{48}$$

In the final step, we just renamed the dummy index of summation. The net result here is that we can swap a lowered index and a raised index for each other in this manner.

Remember how I said above that a tensor was something that existed without a coordinate system? This means that if we write a tensor equation down, that's a relation that the tensors themselves obey, and so *that equation must be true in all coordinate systems*. That is a very powerful statement, and we're going to exploit it a number of times.

Given that, consider what happens if a tensor is zero: $T^\mu{}_{\nu\lambda} = 0$, for example. This could be any arbitrary tensor, so long as it's zero in a particular coordinate system. Note that this is a tensor equation: everything is written in terms of arbitrary spacetime indices, and so we expect the tensor to be vanishing in all coordinate systems. We can check this explicitly – if the tensor vanishes in one coordinate system, then we can transform to any other coordinate system by using a series of Lorentz transformations. Only thing is, they'll all be multiplying zero, and so the net result will be zero. A tensor that is vanishing in any coordinate system is vanishing in all coordinate systems. This fact can be really useful!

Another thing that is really useful is scalar quantities. Because a scalar is the same in any coordinate system, you can calculate a scalar in two separate frames and then compare them. This can be particularly useful in kinematics problems. A scalar quantity is simply any tensor that doesn't have any Lorentz indices on it.

One final bit of terminology before we get back to the physics. When multiplying two tensors together, we will often say that we "contract" the indices. This simply means that we use the Einstein Summation Convention on those indices. For example, consider the quantity $v^\mu x_\nu$. If we contract the indices, we get $v^\mu x_\mu$. This forms a Lorentz scalar, so this quantity will be the same in all reference frames.

I'm going to give you one more bit of advice. If ever you're writing down an equation in terms of this component notation, you will have "free indices" and "contracted indices". Free indices are those that are left over, while contracted indices are those that are summed over with the Einstein Summation Convention. On each side of the equation, you have to have the same free indices left over. If you don't, you know that you've made a mistake somewhere.

7 Momentum

Now that we've defined four-velocity and we know a little bit about tensors, let's define momentum. First of all, we need to define mass. The rest mass of an object is a scalar. It's a Lorentz invariant. We'll denote it m[4]. Momentum is just mass times velocity, so let's stick with that. Here is the momentum four-vector.

$$\begin{aligned} p^\mu &= mv^\mu && (49) \\ &= (m\gamma, m\gamma\mathbf{v}) && (50) \end{aligned}$$

The space components of the four-momentum are the relativistic three-momentum, or $\mathbf{p} = m\gamma\mathbf{v}$, kind of like we would expect. If you remember the expression for relativistic energy, it turns out that the time component is just that. So, we can also write the four-momentum as

$$p^\mu = (E, \mathbf{p})\,. \tag{51}$$

Let's compute p^2.

$$\begin{aligned} p^2 &= p^\mu p^\nu g_{\mu\nu} && (52) \\ &= -E^2 + \mathbf{p}^2 && (53) \end{aligned}$$

But, we already know that $v^2 = -1$, and $p^\mu = mv^\nu$, and so we have $p^2 = -m^2$. This is really important – remember it! Ok, let's put those two formulas together.

$$E^2 = \mathbf{p}^2 + m^2 \tag{54}$$

$$E^2 = \mathbf{p}^2 c^2 + m^2 c^4 \tag{55}$$

In the second line, I reinserted the factors of c, because this formula is pretty iconic. Finally, let's look at what happens if a particle is at rest, with zero three-momentum.

$$E^2 = m^2 c^4 \tag{56}$$

$$E = mc^2 \tag{57}$$

There we go, we've found Einstein's famous equation: the rest energy of a mass is mc^2. Awesome!

What happens if we have a massless particle? We can't use $p^\mu = mv^\mu$ then, because it doesn't make sense. But the four-momentum is still well-defined; we still have

$$p^\mu = (E, \mathbf{p})\,. \tag{58}$$

Because $p^2 = 0$, we then have $E = |\mathbf{p}|$, and a massless particle's energy is simply determined from it's momentum. Now, remember from our definition of the four-momentum, we said that we had $E = m\gamma$, and $\mathbf{p} = m\gamma\mathbf{v}$? If we put these two together, we get $\mathbf{p} = E\mathbf{v}$, which holds even as we take the limit of zero mass. This tells us that $|\mathbf{p}| = E|\mathbf{v}|$. Substituting this into the relationship between energy and momentum for a massless particle, we find that $|\mathbf{v}| = c$. A massless particle has to travel at the speed of light!

Now that we have a four-momentum, what's it good for? Well, the four-momentum tracks the energy and the momentum of a particle. If we have multiple particles colliding, they'll need to conserve energy and momentum. If we conserve four-momentum, we do both in one go!

[4] Some authors call it m_0, and reserve m for the "relativistic mass". I don't like this terminology, because it suggests that mass increases with velocity, which is a questionable statement at best.

Ok, so we have four vectors for position, velocity, and momentum. Can we keep on going? What about a four-acceleration? Perhaps a four-force? As you might expect, you can define them. Are they useful? Not really, to be honest. I've derived four-acceleration before, but it's a bit of a mess. A four-force is even worse. And you know what? I've never used them in my life. So, we'll stop there, and not worry about them. We will actually come to a four-force a little later on, but we'll leave it for now[5].

8 Fields

Now that we've seen how to write down stuff for vectors and point particles and so on, we need to talk about fields. What we would like to do is to create a field at every point in spacetime, similarly to how we do so in Newtonian mechanics, where we write a field as $\phi(t, \mathbf{x})$. Vector fields, such as the electric field, will become a little more tricky, as we will see shortly. In this section, we'd like to discuss fields, and how they transform under Lorentz transformations.

Let's start with a scalar field. This is a field that takes on a value at every point in spacetime. We'll call our hypothetical scalar field $\phi(x^\mu)$. This could be something like the number of apples at each spacetime location. Different observers will all agree on the the number of apples. However, the labels that different observers put on that spacetime location will vary. In particular, we have

$$\phi(x^\mu) = \phi'\left(\Lambda^{\mu'}{}_{\mu} x^\mu\right). \tag{59}$$

Because what I see as x^μ is what you see as $\Lambda^{\mu'}{}_{\mu} x^\mu$, we need to update our book-keeping to get from ϕ to ϕ'.

The next field to consider is a vector field. Examples could be a velocity field, a momentum field, or as we shall later see, the electromagnetic potential field. For these, not only do we need to update what label we use for the event that we're looking at, we also need to update what direction the field is pointing in. How do we do that? We use a Lorentz transformation, of course!

$$v^\mu(x^\nu) = \Lambda^{\mu}{}_{\mu'} v^{\mu'}\left(\Lambda^{\nu'}{}_{\nu} x^\nu\right) \tag{60}$$

Note that we have to use a Lorentz transformation to update the $v^{\mu'}$ part, while we have to use the inverse Lorentz transformation to update the x^ν part. This is tricky. Go and read that again, and make sure you understand why it's happening (look back at how the scalar field transforms if you need a hint). Note that if we look at the free indices on both sides of the equation, they're the same. We have different indices on the arguments, but that's ok.

We can get more complicated than a vector field, by introducing fields that have more than one index. We'll see one of these – the electromagnetic field strength tensor. It looks like $F_{\mu\nu}$. Can you guess how this will transform under a Lorentz transformation? We'll need one Λ for each index, and we'll also need to transform the coordinates that we're looking at.

$$F_{\mu\nu}(x^\lambda) = \Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu} F_{\mu'\nu'}\left(\Lambda^{\lambda'}{}_{\lambda} x^\lambda\right) \tag{61}$$

Does that make sense? Understanding how to transform this tensor is going to be the *most important thing you learn from this part of the course.* Yes, that's right, if you get nothing else out of these notes, make sure you know how to transform the tensor $F_{\mu\nu}$.

8.1 Derivatives of Fields

Now that we have some definitions of fields, we're going to want to differentiate them. Let's start with a scalar field, $\phi(x^\mu)$. Say I wanted to know the gradient in the x direction. That would look something like

[5] I'm not exactly sure why, but it is a common misconception that you cannot deal with accelerations in special relativity. Special relativity can treat accelerations just fine, although the math tends to become ugly.

the following.

$$\frac{\partial \phi(x^\mu)}{\partial x} \tag{62}$$

A better way to write that would be as the partial derivative with respect to x^1, which is the x coordinate.

$$\frac{\partial \phi(x^\mu)}{\partial x^1} \tag{63}$$

Even better than that, we could let it be a derivative in an arbitrary direction:

$$\frac{\partial \phi(x^\mu)}{\partial x^\nu}. \tag{64}$$

This will give us a function with one spacetime index, ν. Should that index be raised or lowered? Let's work it out in an example. Let $\phi(x^\mu) = x^\mu v_\mu$ where v_μ is a constant vector. Then we have

$$\frac{\partial \phi(x^\mu)}{\partial x^\nu} = \frac{\partial (x^\mu v_\mu)}{\partial x^\nu} = v_\mu \frac{\partial x^\mu}{\partial x^\nu}. \tag{65}$$

What is this last strange beast? If you work it out component by component, you should find that it gives δ^μ_ν. Thus, our result here would be v_ν. So, we started with a scalar object, and ended up with a vector object with a lowered index. Differentiating a field with respect to a spacetime coordinate yields a new field with an extra lowered index. This is important. Read it again.

In fact, it's so important, that we give a shortcut to the notation. If we want to take a derivative with respect to a coordinate, we often write the following:

$$\frac{\partial}{\partial x^\mu} \equiv \partial_\mu. \tag{66}$$

This makes it really obvious that we're adding an extra spacetime index. Can you guess how to Lorentz transform the derivative of a field? We need to act on the new spacetime index with another Lorentz transformation. Exercise: Write out the complete Lorentz transformation of the partial derivative of a scalar field. (Hint: It should be exactly the same as for a vector field with a lowered index!)

There are a couple of subtleties we need to be careful about. The first is the following. What would it mean to write ∂^μ? This always means $g^{\mu\nu}\partial_\nu$, and you should always evaluate it as such. Don't think about it as a derivative with regards to a lowered position vector. It will get you confused.

The second subtlety to think about is the following. What is the derivative $\partial_\mu x_\nu$? Take a moment to think about how you'd evaluate this. What you should be thinking is that you need to raise the index on x before you can differentiate, something like this:

$$\partial_\mu x_\nu = \partial_\mu \left(g_{\nu\lambda} x^\lambda\right) = g_{\nu\lambda} \partial_\mu x^\lambda = g_{\nu\lambda} \delta^\lambda_\nu = g_{\mu\nu} \tag{67}$$

Note that this works because the metric is not position dependent. So, if you come across a field that is written with lowered position variables, you'll need to raise the indices with the metric appropriately to take derivatives. If you think you can do this, try the following exercise: What is $\partial_\mu(x^\nu x_\nu)$?

9 Interlude

You are now armed with all of the knowledge necessary to do absolutely everything in special relativity. At least, you've seen it, and if you continue with physics, you'll see a lot more of it, and you'll get a lot more comfortable with it. If you've just read to this point, you've done all of the hard stuff, and all we're going to do now is apply what you've learnt here to electromagnetism. Take a deep breath, and let's dive in and have some fun!

10 Electrostatics

We've seen that if we write things down as 4-vectors, then we can use Lorentz transformations to change coordinate systems really easily. It works for position, velocity, and energy-momentum (usually just called momentum). Can we do the same thing for the electric field?

The key to constructing the four-vector was to identify an appropriate fourth component. To go with space, we added time. To go with velocity, we added a normalization condition. To go with momentum, we added energy. What can go with an electric field, which normally has three components? We need something that would transform into the electric field and vice-versa under a Lorentz transformation.

We're at a bit of a loss, alas. It turns out that there is no good four-vector to describe the electric field. What we instead want to look at is just the electric potential. Is it just a scalar field? Well, sort of. In three dimensions, it's a scalar field, and it doesn't transform under rotations. When you consider rotations in spacetime though, it many not still be a scalar field – it could be the time component of a four-vector, for example, and it turns out that that's what we want. Let's try the following, and write the electric potential as the first component in a new four-vector. We'll call it[6] $A^\mu = (\phi, \mathbf{0})$. Given that we don't know what to put for the vector yet, let's assume that it's vanishing. Let's assume that we have a situation where all of our charges are stationary. We know how everything behaves in electrostatics, so that motivates that anything that's in the spatial components is vanishing.

So we've defined a four-vector. What next? Well, what we really want to do is relate the electric field to the potential. Consider the following tensor, which I am just going to pull out of the air. This is the electromagnetic field strength tensor (and you'll see why very shortly).

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{68}$$

If you calculate the four components of this tensor, you should find that $F_{0i} = -E_i$, where E_i is the component of the electric field in the i direction (remember to use $\mathbf{E} = -\nabla\phi$, and also remember to lower the index on A^μ). Note that the electric field is a 3-vector, so it doesn't really matter if the index is raised or lowered (remember that the metric in three dimensions is just $(1, 1, 1)$ on the diagonals). You should also see that $F_{i0} = E_i$. By looking at the definition of $F_{\mu\nu}$, convince yourself that this makes sense, because F is antisymmetric in μ and ν by definition.

That's kind of cute. By writing the potential as the time component of a four-vector, we get the electric field as the components in a tensor. What about the other components of the tensor? F is vanishing for all of them. There's nothing there.

Now, let's see what our favorite equations look like in this form. What favorite equations, you say? Why, these ones, of course!

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} = 4\pi\rho \tag{69}$$

$$\nabla \times \mathbf{E} = 0 \tag{70}$$

Look familiar? They should! Recall that both of these equations come simply from Coulomb's Law. For the purposes of special relativity, cgs units are just easier to work with, and so we'll use that convention here. Remember that we're using units where $c = 1$.

Let's start with the second of these first. I'm going to claim that the following equation is the one that we want:

$$\partial_\gamma F_{\mu\nu} + \partial_\mu F_{\nu\gamma} + \partial_\nu F_{\gamma\mu} = 0 \tag{71}$$

There are a lot of equations here – we have three free indices, so it naively looks like there are 64(!) equations to check. Thankfully, there's not really that many, and in fact, it is possible to show that there are only four

[6] The assumption that we should write the electric potential as part of a four-vector is an important one. In physics lingo, we say that the electromagnetic field is a spin one field. That is, the electromagnetic potential has one Lorentz index. If you try to do the same thing for gravity, which looks like it behaves exactly the same as electrostatics, you'll find that you end up with the wrong answers. This is because gravity is not a spin one field; gravity is a spin two field, and that is the subject of general relativity.

independent equations. Let's look at the equation for $\gamma = x$, $\mu = y$, and $\nu = t$.

$$0 = \partial_x F_{yt} + \partial_y F_{tx} + \partial_t F_{xy} \tag{72}$$
$$= \partial_x E_y - \partial_y E_x \tag{73}$$

This should look vaguely familiar. In fact, it's the $\hat{z}$ component of $\nabla \times \mathbf{E} = 0$! Exercise: Find the other two equations that give the $\hat{x}$ and $\hat{y}$ components of $\nabla \times \mathbf{E}$.

You might be thinking "That's three equations – what's the fourth?". Hold onto that thought for now, we'll come back to it. For the moment, that fourth equation just says $0 = 0$, which is kind of meaningless.

Ok, so that's one of our equations of electrostatics. What about the other? Well, we're going to need to introduce the charge density, ρ. What happens if we perform a Lorentz transformation on a charge density? Well, if we just do a rotation, we'll not really change anything – the charge is still in the same place, etc. But if we perform a Lorentz boost, then we'll get a current. So, current and charge should be related in some way. The current will have a direction, so that should play the role of the spatial component, while the charge density is directionless, so it can take the time component. Let's define the current four-vector $J^\mu = (\rho, \mathbf{0})$ for the moment. We'll get to currents soon.

I'm going to once again pull an equation out of thin air:

$$\partial_\mu F^{\nu\mu} = 4\pi J^\nu. \tag{74}$$

Be careful, fair reader. I have invoked a couple of raised metrics here. (Also be careful on the order of the indices!) There are four equations here, one for each ν. If ν is spatial (x, y, or z), the right hand side vanishes, and so we get a strange equation (by combining all three equations):

$$\frac{\partial \mathbf{E}}{\partial t} = 0 \tag{75}$$

This tells us that the electric field is not varying in time. That's what we expect though, given that we assumed that we were talking about electrostatics.

If we look at the time component of this equation though, we get something exciting:

$$\partial_x E_x + \partial_y E_y + \partial_z E_z = 4\pi\rho \tag{76}$$

This is our old friend, $\nabla \cdot \mathbf{E} = 4\pi\rho$.

So, we've done something really quite cute – by writing things in this four-vector notation, we've captured the equations of electrostatics in one neat bundle. The next step is to look at what happens when we Lorentz transform!

11 Lorentz Transformed Electrostatics

Let's look at what happens if we Lorentz transform the electromagnetic field strength tensor. Recall the formula from above to do so:

$$F_{\mu\nu}(x^\lambda) = \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu F_{\mu'\nu'} \left(\Lambda^{\lambda'}{}_\lambda x^\lambda \right) \tag{77}$$

For the moment, I'm going to assume that we have a constant electric field in all of space, so we can ignore the spacetime dependence. Next, because we have a constant electric field, we might as well rotate our axes so that it's pointing in the x direction. Here is what our electromagnetic field strength tensor looks like (first index row, second index column):

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & 0 & 0 \\ E_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{78}$$

What do we get when we transform it? Let's boost in the y direction, say. Go and do this transformation. No, really. Go and do it. You can calculate each component individually, or a trick that works here because so few quantities are nonzero is to write out the following:

$$F_{\mu'\nu'} = \Lambda^0{}_{\mu'}\Lambda^1{}_{\nu'}F_{01} + \Lambda^1{}_{\mu'}\Lambda^0{}_{\nu'}F_{10} \tag{79}$$

Evidently, the $\Lambda^1{}_{\nu'}$ and similar coefficients will only be nonzero for $\nu' = 1$, so we can quickly figure out what the nonzero terms actually are. Here they are in all their glory:

$$F_{\mu\nu} = \begin{pmatrix} 0 & -\gamma E_x & 0 & 0 \\ \gamma E_x & 0 & -v\gamma E_x & 0 \\ 0 & v\gamma E_x & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{80}$$

As usual, $\gamma = 1/\sqrt{1-v^2}$. Remember to use the inverse Lorentz transformation!

Ok, this is a little unusual. Our electric field is still there, but it's showing up in another place as well. From the perspective of the new reference frame, that quantity v is completely meaningless (it just relates two different frames, nothing more), and so it appears that there's a funny new quantity showing up. Can this just be a redundant description of the electric field? No, because of the presence of v in the expression. Thus, we actually have a strange and mysterious new field now appearing. It only has one component present at the moment, but you can easily imagine that if instead of boosting in the y direction, we boosted in the z direction, we'd get a different component turning up. Similarly, if we start with the electric field in the y direction, we'd get a third component showing up. Given that there are three components, for the moment, we'll call it the **B** field, and you can probably guess that it will turn out to be the magnetic field. Where did it come from? It appeared because the electric field needed something to transform into! It couldn't just transform into itself; it had to transform into something else as well. In the next section, we'll figure out what the equations of motion for this new field are.

12 Electromagnetism: The Whole Kit, Kat and Kaboodle

Let's go back to our formulation of electrostatics. We started with the electric potential, ϕ, in a four-vector, A^μ. Now that we've seen that a Lorentz transformation turns the potential into the other components of the four-vector, let's give those components a name. We'll call them $\mathbf{A}$, the magnetic vector potential. The four-vector will then be $A^\mu = (\phi, \mathbf{A})$. You can probably guess that it's related to this new mysterious field. We already identified above that Lorentz transforming the charge density would yield a current density in the other slots of the four-vector, so $J^\mu = (\rho, \mathbf{J})$, where $\mathbf{J}$ is the usual three-vector current density.

Now, assume that we start with some electrostatic situation, and we Lorentz transform it to a moving frame. We no longer expect the situation to be static, because time and space are mixed. Furthermore, we know that we now have currents, and our four-vector potential has developed spatial components too. Now for the most important part of all: do you remember what we said about tensor equations above? Those three equations that we've written down, relating the four-potential to the field strength, and giving our curl and divergence of $\mathbf{E}$, are all tensor equations. Thus, we expect them to be true, regardless of the coordinate system that we use. Said another way, we want physics to agree in all reference frames. So, what we know occurs in the static frame must also occur in this moving frame, and thus the Lorentz transformation of the equations in the static frame must be valid in the new frame. Read that again. It's a very powerful statement.

Instead of constructing the transformation from some original frame, let's assume that we have a four-vector potential, and a four-current. We know what the equations describing everything were in the original frame; they must also be true in the new frame. So, let's go and find out what those equations have to say in the new frame.

When we differentiate the four-vector potential to obtain the field strength tensor, we find the following. Our old friend, the electric field, has something new to tell us:

$$F_{i0} = E_i = -\partial_i\phi - \partial_0 A_i \tag{81}$$

In vector notation, this tells us that

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}. \tag{82}$$

So, once things get time dependent, the rate of change of that extra vector in the potential becomes important. Interesting. What about the other components? Well, we have three nonzero equations, and they are as follows.

$$F_{12} = \partial_1 A_2 - \partial_2 A_1 \tag{83}$$
$$F_{23} = \partial_2 A_3 - \partial_3 A_2 \tag{84}$$
$$F_{31} = \partial_3 A_1 - \partial_1 A_3 \tag{85}$$

The antisymmetric partners of these hold no new information. What this really looks like is a cross product, the way those components are paired up. So, let's call our new field, $\mathbf{B} = \nabla \times \mathbf{A}$, such that $B_x = F_{23}$, $B_y = F_{31}$, and $B_z = F_{12}$, and just because we can, we'll call it our magnetic field. Our electromagnetic field strength tensor then takes the following form (again, first index row, second index column):

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{86}$$

Now this looks complete.

Next, let's look at the equations that gave us our electrostatic equations waaay back above. The first one, which told us that the curl of $\mathbf{E}$ was vanishing, has mysteriously changed.

$$\partial_\gamma F_{\mu\nu} + \partial_\mu F_{\nu\gamma} + \partial_\nu F_{\gamma\mu} = 0 \tag{87}$$

Look at the $\gamma = x$, $\mu = y$, and $\nu = t$ component of this equation.

$$0 = \partial_x F_{yt} + \partial_y F_{tx} + \partial_t F_{xy} \tag{88}$$
$$= \partial_x E_y - \partial_y E_x + \partial_t B_z \tag{89}$$

Huh... it appears that the curl of the electric field has something to do with the time derivative of the magnetic field. Exercise: Find the other two equations that go with this. Combining these three equations, we find the following for the curl of $\mathbf{E}$.

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{90}$$

Now, remember how I said there was a fourth equation here that was previously trivial? Now it gives us some information. Look at $\gamma = x$, $\mu = y$, $\nu = z$.

$$\partial_x F_{yz} + \partial_y F_{zx} + \partial_z F_{xy} = 0 \tag{91}$$
$$\partial_x B_x + \partial_y B_y + \partial_z B_z = 0 \tag{92}$$
$$\nabla \cdot \mathbf{B} = 0 \tag{93}$$

This new equation tells us that the magnetic field is divergenceless.

The other equation that told us about the divergence of the electric field has new information too.

$$\partial_\mu F^{\nu\mu} = 4\pi J^\nu. \tag{94}$$

Our old equation, the $\nu = 0$ component, has nothing new to add. We still have $\nabla \cdot \mathbf{E} = 4\pi\rho$. On the other hand, the other three equations have exciting and new information for us! Looking at the $\nu = x$ component, for example, we find

$$\partial_t F^{xt} + \partial_x F^{xx} + \partial_y F^{xy} + \partial_z F^{xz} = 4\pi J^x \tag{95}$$
$$-\partial_t E_x + \partial_y B_z - \partial_z B_y = 4\pi J^x \tag{96}$$

You can work out the y and z components. Exercise: Show that these three equations can be combined to give the following,

$$\nabla \times \mathbf{B} = 4\pi \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t} \tag{97}$$

where $\mathbf{j}$ is the three-vector current density.

We thus have four equations governing how $\mathbf{E}$ and $\mathbf{B}$ work. These are known as Maxwell's equations. I'll repeat them again here, because they're really important.

$$\nabla \cdot \mathbf{E} = 4\pi\rho \tag{98}$$

$$\nabla \times \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} \tag{99}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{100}$$

$$\nabla \times \mathbf{B} = 4\pi \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t} \tag{101}$$

Also, we have a way of writing our fields in terms of potentials. The equations for the potentials are:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{102}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{103}$$

I will note that in the process of restoring units here, we end up finding a new constant of nature, which we call μ_0, which is important in the magnetic equations. This constant is related to ϵ_0 and the speed of light in the following way:

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \tag{104}$$

It took Maxwell's combined equations and a lot of work to discover that those three quantities were all related, and from our above derivation, we found that we never needed to introduce one in the first place. Neat, huh? Exercise: Go and figure out how to insert ϵ_0 and μ_0 appropriately into Maxwell's equations using dimensional analysis. You shouldn't have any factors of c left over.

Let's stop here and admire those equations a bit. Just from Coulomb's Law and Special Relativity, we've derived all of those equations, which amounts to almost all of electromagnetism. There's a little bit more that we'll see in the section below on the Lorentz Force Law, and that will make everything. You're going to spend the rest of the course investigating what each of these equations mean, and how all of the terms in each of these equations behave.

13 Transforming Electromagnetism

We now have all of electromagnetism written down in our nice component notation. Furthermore, we know how to Lorentz transform these quantities. Note that all of the equations that we are playing with are *linear* in the electric and magnetic fields (or alternatively, the electromagnetic four-potential). This means that we can understand how electric fields transform by themselves, and then how magnetic fields transform by themselves, and get the whole picture by superimposing the two. Furthermore, because of this linearity, you can see that like the electric field, the magnetic field also obeys the superposition principle. To find the general transformation law is a bit of a pain. If you'd like to see what it looks like, take a look at `http://en.wikipedia.org/wiki/Lorentz_transformation`. It is possible to apply these formulas blindly, but the process I have developed here highlights the relativistic nature of these fields, and is much, much easier to remember.

Knowing how to transform the electric and magnetic fields is the most important skill to developed from this part of the course. A number of examples of doing this are part of the problem set at the end of these notes, and I encourage you to practise them. Here, I want to look at an important example that is also mathematically nontrivial, so as to show you some interesting physics, and to spare you the pain of doing it yourself on the homework.

13.1 Example: A Moving Charge

We know what a point charge looks like in electrostatics. What we would like to do is to understand how the field behaves when we look at it from the perspective of a moving observer.

To start, let's write down the static configuration.

$$A^\mu = \left(\frac{Q}{r}, \mathbf{0}\right) \tag{105}$$

$$\mathbf{E} = \frac{Q}{r^2}\hat{r} \tag{106}$$

$$\mathbf{B} = 0 \tag{107}$$

$$J^\mu = Q(\delta^3(\mathbf{x}), \mathbf{0}) \tag{108}$$

The coordinate r is a bit of a pain to work with in special relativity, because everything is done in Cartesian coordinates. So, you'll have to mentally substitute in $r = \sqrt{x^2 + y^2 + z^2}$. Note that our charge is a delta function located at the origin.

Let's boost into a moving frame now, to the perspective of someone moving in the positive x direction with respect to us. The four-potential becomes

$$A^{\mu'} = \left(\frac{Q}{r'}\gamma, -\frac{Q}{r'}v\gamma, 0, 0\right) \tag{109}$$

where $r' = \sqrt{(\gamma x' + v\gamma t')^2 + y'^2 + z'^2}$. Remember how we need to inverse Lorentz transform the argument of our functions? That's why r had to transform.

We'll do the other easy one next, the four-current. In the new frame, this vector is

$$J^{\mu'} = Q(\gamma\delta(\gamma x' - v\gamma t')\delta(y')\delta(z'), -v\gamma\delta(\gamma x' - v\gamma t')\delta(y')\delta(z'), 0, 0). \tag{110}$$

Now, one property of the delta function is that for constant positive a, $a\delta(x) = \delta(x/a)$, and so we can rewrite our four-current as

$$J^{\mu'} = Q(\delta(x' + vt')\delta(y')\delta(z'), -v\delta(x' + vt')\delta(y')\delta(z'), 0, 0). \tag{111}$$

This is exactly what we expect for our current; it looks like the same charge Q, traveling backwards along the x axis.

To calculate the electric and magnetic fields, we can either transform $F_{\mu\nu}$, or we can differentiate the new four-potential. If you have the potential, typically the latter will be simpler. In terms of our usual variables, we have

$$\phi = Q\gamma\frac{1}{\sqrt{\gamma^2(x + vt)^2 + y^2 + z^2}} \tag{112}$$

$$\mathbf{A} = -Qv\gamma\frac{\hat{x}}{\sqrt{\gamma^2(x + vt)^2 + y^2 + z^2}} \tag{113}$$

where I'm dropping the primes for simplicity. Now that our special relativity is done, it's time to find a more convenient choice of coordinates. Let's use cylindrical coordinates, which are (I think) best suited towards this calculation. Let $y^2 + z^2 = \rho^2$, and $x = h$ (note that this isn't the usual transformation to a cylinder lying along the z axis).

$$\phi(t, h, \rho) = Q\gamma\frac{1}{\sqrt{\gamma^2(h + vt)^2 + \rho^2}} \tag{114}$$

$$\mathbf{A}(t, h, \rho) = -Qv\gamma\frac{\hat{h}}{\sqrt{\gamma^2(h + vt)^2 + \rho^2}} \tag{115}$$

Now, we can calculate our fields.

$$
\begin{aligned}
\mathbf{E} =& -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} && (116)\\
=& -\frac{\partial \phi}{\partial h}\hat{h} - \frac{\partial \phi}{\partial \rho}\hat{\rho} - \frac{\partial \mathbf{A}}{\partial t} && (117)\\
=& \, Q\gamma \left(\hat{h}(h+vt) + \hat{\rho}\rho\right) \frac{1}{\left(\gamma^2(h+vt)^2+\rho^2\right)^{3/2}} && (118)
\end{aligned}
$$

The (not too) messy algebra is left as an exercise for the reader (it helps to note that $\mathbf{A} = -v\phi\hat{h}$, and that $\partial\phi/\partial t = v\partial\phi/\partial h$). To understand what this equation means, it's probably worthwhile to look at our original electric field in cylindrical coordinates.

$$
\mathbf{E}_0 = Q\frac{\rho\hat{\rho} + h\hat{h}}{\left(h^2+\rho^2\right)^{3/2}} \tag{119}
$$

There are three main differences. The first is that you are now moving with respect to the charge, so the distance $h \to h + vt$, which was to be expected. The second is that the field is strengthened by a factor of γ. The third, and most important effect, is that the $(h+vt)$ in the calculation of distance in the denominator is boosted by γ^2. This means that distances along the direction of motion are much more stretched – they've gone from Δx away from the charge to effectively $\Delta x/(1-v^2)$, which at high speeds, can be much, much larger. This implies that the electric field about the point charge is much stronger in the plane perpendicular to it's motion than in the direction parallel to its motion. Figure 5.14 in Purcell describes this.

Let's look at the magnetic field. This calculation in cylindrical coordinates is actually relatively straightforward.

$$
\begin{aligned}
\mathbf{B} =& \nabla \times \mathbf{A} && (120)\\
=& \, Qv\gamma \frac{\partial}{\partial h}\frac{1}{\sqrt{\gamma^2(h+vt)^2+\rho^2}}\hat{\phi} && (121)\\
=& -Qv\gamma \frac{(h+vt)\gamma^2}{\left(\gamma^2(h+vt)^2+\rho^2\right)^{3/2}}\hat{\phi} && (122)
\end{aligned}
$$

This is a bit of a weird one. It's only in the $\hat{\phi}$ direction, so it's a winding field. If $h > -vt$, we're looking in the charge's "wake", and the field curls one way. If $h < -vt$, then we're "in front" of the charge, and we feel it's approach. Right on top of the charge, $h = -vt$, it vanishes. At the same time, the factor in the denominator is suppressing anything too far away in the x direction, so the magnetic field will be fairly localized around the charge.

14 The Lorentz Force Law

The last thing that we want to understand is "How does a charge move in electric and magnetic fields?" We already know part of the answer, from the Coulomb force law. Indeed, we're going to start with that, and perform a Lorentz transformation to discover the Lorentz force law.

What we start with is the Newtonian force law that we all know and love:

$$
\mathbf{F} = \frac{d\mathbf{p}}{dt} = q\mathbf{E}. \tag{123}
$$

The first thing we do is to talk about the change in momentum rather than the force, because that will make more sense to us later on than an abstract force. Next, we want to construct a compoment notation equation for this. $\mathbf{E} = F^{0i}$ is straightforward, but we'll eventually want to make a spacetime statement rather than just a statement about these particular coordinates. Let's look towards momentum. The quantity $d\mathbf{p}/dt$ is difficult, because that derivative is problematic, as we discussed before. What's better is to look at $d\mathbf{p}/d\tau$, the derivative with respect to proper time. We'll need to introduce a factor of γ to do so, just like when we defined four-velocity. However, we're looking at our particle at rest, so $\gamma = 1$. So, let's then write

$$\frac{dp^i}{d\tau} = qF^{0i}. \tag{124}$$

Ok, that's got the shape of the kind of equation we want to play with. We need to do two things. The first is to somehow turn that zero into a proper spacetime index rather than something that doesn't transform, and the second is to extend the i index to t as well. Let's tackle the zero first. Exercise: Explain why having a time component here doesn't lead to a good equation in special relativity.

The trick is to note that we are looking at electrostatics, so we've got objects at rest. When a particle is at rest, it's velocity is $v^\mu = (1, 0, 0, 0)$, and so we can use this to pick out the zeroth component of the electromagnetic field strength tensor. Let's put that in, and note that we also need to include a minus sign, because of the metric (here, I've absorbed the minus sign by flipping the order of the indices on F).

$$\frac{dp^i}{d\tau} = qv_\mu F^{i\mu}. \tag{125}$$

Ok, that looks like a good equation. Now, how about extending the equation to all spacetime indices? Let's look at what the time component of this equation would say.

$$\frac{dE}{d\tau} = qv_\mu F^{0\mu} = 0. \tag{126}$$

Exercise: Show that the right hand side is zero. You'll want to use what you know about v_μ in the rest frame, and one of the properties of $F^{\mu\nu}$.

That's a little strange. If our particle is about to be pushed by the electric field, is the first derivative of energy vanishing? Let's check, by Taylor expanding energy.

$$p^0 = E = \sqrt{m^2 + \mathbf{p}^2} = \sqrt{m^2 + \gamma^2 m^2 \mathbf{v}^2} \tag{127}$$

$$= m + \frac{m}{2}\mathbf{v}^2 + O(\mathbf{v}^4) \tag{128}$$

$$\frac{dE}{d\tau} = \frac{dm}{d\tau} + \frac{m}{2}\frac{d\mathbf{v}^2}{d\tau} + O(\mathbf{v}^4) \tag{129}$$

$$= 0 + m\mathbf{v} \cdot \frac{d\mathbf{v}}{d\tau} + O(\mathbf{v}^4) \tag{130}$$

$$= 0 \tag{131}$$

Indeed, it does vanish, because $\mathbf{v} = 0$. The second derivative wouldn't be zero though, so something is happening. Exercise: Show that if you use the fully relativistic expression for energy, the first derivative is still vanishing.

So, it appears that we can happily include the time component of our equation, and it's a valid equation describing our static case. Next, we apply the same argument as before, and Lorentz transform, claiming that physics must be the same to any other observer, and so our equation must again be valid. So, here's our force law in a different frame of reference:

$$\frac{dp^\nu}{d\tau} = qv_\mu F^{\nu\mu} \tag{132}$$

Let's look at the four separate components here. The time component we just looked at before. For an arbitrary moving charged body, it gives us

$$\frac{dE}{d\tau} = q\gamma \mathbf{E} \cdot \mathbf{v}. \tag{133}$$

Exercise: derive this from the previous equation. Let's integrate this equation to see what it says.

$$\Delta E = q \int \mathbf{E} \cdot \frac{d\mathbf{x}}{dt} dt = q \int \mathbf{E} \cdot \mathbf{dx} = -q\Delta\phi \tag{134}$$

That's exactly what we want, from our knowledge of electrostatics. Furthermore however, this is general – we have magnetic fields here too. What this tells us is that magnetic fields cannot change the energy of a particle.

Ok, that's the time component, let's look at the space component, which we expect to give us something like $\mathbf{F} = q\mathbf{E}$. The x component yields

$$\frac{dp^x}{d\tau} = q\gamma \left(E^x + v^y B^z - v^z B^y\right). \tag{135}$$

That looks suspiciously like a cross product. Exercise: Show that the remaining two equations yield the following vector equation:

$$\frac{d\mathbf{p}}{d\tau} = \gamma q \left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{136}$$

If we now convert this derivative in terms of proper time into coordinate time, we arrive at the famous Lorentz Force Law:

$$\frac{d\mathbf{p}}{dt} = q \left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{137}$$

Note that this is valid for any velocity $\mathbf{v}$!

15 Conclusion

And there we go – that's everything. We have now derived all of electromagnetism, starting from Coulomb's Law and Special Relativity. Pretty impressive, isn't it? In the rest of the course, you'll be looking at Maxwell's equations in detail, term by term, and understanding how they work. I hope you've enjoyed this little romp, which is one of my favorite pieces of physics this universe has to offer.

16 Appendix: Even More Interesting Stuff! (Bonus material)

You may recall that I said that scalars are really useful things in special relativity. However, I didn't calculate any scalars pertaining to the electromagnetic field. Here, we'll rectify that little issue. One scalar we can construct is $A_\mu A^\mu$. However, this isn't really all that physical, and it doesn't contain any information, so we'll ignore it. The next one is $F_{\mu\nu}F^{\mu\nu}$, which is really useful. Exercise: Show that this scalar is given by the following.

$$F_{\mu\nu}F^{\mu\nu} = 2(\mathbf{B}^2 - \mathbf{E}^2) \tag{138}$$

Apart from the metric and the Kronecker delta, there turns out to be one more tensor object that is the same in every coordinate system. This is the completely antisymmetric tensor (also known as the Levi-Civita tensor, or the epsilon tensor), ϵ^{abcd}. This object is defined by

$$\epsilon^{0123} = +1 \tag{139}$$

where every swap of two indices swaps the sign. For example, $\epsilon^{0132} = -1$ because of one index swap, while $\epsilon^{0312} = +1$ because of two index swaps. If you lower the indices using the metric, you will find that you get $\epsilon_{0123} = -1$. (Sometimes, it is defined as $\epsilon_{0123} = +1$ instead, so make sure which definition you're using.) If you Lorentz transform this object, you'll find that it is the same in every coordinate system. (Try this, if you like. You'll need to figure out what the determinant of a Lorentz transformation metric is. As a hint, remember that Lorentz transformations as we constructed them are in the $SO(3,1)$ rotation group. The O stands for orthogonal matrices, which is a bit tricky with the funny signature, and the S stands for "special". See if you can find out what "special" means.)

What's it useful for? Let's look at the three-dimensional version. In three dimensions, the epsilon tensor is $\epsilon^{123} = +1$. (Because the metric in three dimensional space is Euclidean, we can lower the indices and get $\epsilon_{123} = +1$, so we don't worry about raised and lowered indices in Euclidean space.) Now, consider the following expression.

$$(\mathbf{r} \times \mathbf{p})^i = \epsilon^{ijk} r^j p^k \tag{140}$$

The epsilon tensor is really useful for writing down cross products! Kinda neat. (Ignore the dodgy summation convention usage here; it comes about because indices can be raised and lowered at will in three dimensions.)

What's the four-dimensional version good for? Do you remember how we were trying to make Lorentz invariant things out of the electromagnetic field tensor, F_{ab}? We found this one:

$$F_{\mu\nu} F^{\mu\nu} = 2(\mathbf{B}^2 - \mathbf{E}^2). \tag{141}$$

Using the epsilon tensor, we can also construct the following Lorentz invariant quantity.

$$\epsilon^{\mu\nu\sigma\lambda} F_{\mu\nu} F_{\sigma\lambda} = -8\mathbf{E} \cdot \mathbf{B}. \tag{142}$$

So, regardless of what frame you're looking in, at any point, both $\mathbf{E}^2 - \mathbf{B}^2$ and $\mathbf{E} \cdot \mathbf{B}$ must be the same.

That's kind of useful. For example, we've learnt that the magnetic field is just a Lorentz transformation of the electric field. There's probably some condition that says when we can find a frame which only has an electric field or only has a magnetic field (at a point, at least). Let's find it. If $\mathbf{E} \cdot \mathbf{B} \neq 0$, then we *must* have both $\mathbf{E}$ and $\mathbf{B}$ nonzero in any reference frame. So, if we want a frame with only one of the two fields, we will require $\mathbf{E} \cdot \mathbf{B} = 0$. Now, let's look at the other scalar. If $\mathbf{E}^2 - \mathbf{B}^2 > 0$, then we should be able to find a frame that only has an electric field (as $\mathbf{E}^2 \geq 0$). Similarly, if $\mathbf{E}^2 - \mathbf{B}^2 < 0$, we should be able to find a frame with only a magnetic field. What about if $\mathbf{E}^2 - \mathbf{B}^2 = 0$? Well, there's the trivial solution of "there are no fields". Is there a nontrivial solution? Let's pretend that we can find a frame with only a magnetic field. Then we have

$$\mathbf{E}_1^2 - \mathbf{B}_1^2 = 0 = -\mathbf{B}_2^2. \tag{143}$$

Thus, we need $\mathbf{B}$ to vanish in this frame too. The same argument applies for the electric field. So, we've found the condition on finding a frame with only one of the two fields:

$$\mathbf{E} \cdot \mathbf{B} = 0 \tag{144}$$

$$\mathbf{E}^2 - \mathbf{B}^2 \neq 0 \tag{145}$$

This second condition is rather important: in the case of an electromagnetic wave, you will later see that $|\mathbf{E}| = |\mathbf{B}|$, as well as $\mathbf{E} \cdot \mathbf{B} = 0$. This implies that in no frame does an electromagnetic wave look like just a magnetic field or just an electric field.

Here's something else we can use the epsilon tensor for. Recall this equation:

$$\partial_\gamma F_{\mu\nu} + \partial_\mu F_{\nu\gamma} + \partial_\nu F_{\gamma\mu} = 0 \tag{146}$$

Remember how I said there are only four actual equations here? Well, here's how we can show that. Firstly, check the following. If we interchange any two coordinates on the left hand side, every term will pick up a negative sign (you will need to use the fact that $F_{\mu\nu}$ is antisymmetric in its indices, and also to swap a couple of terms around). This means that the entire expression on the left hand side is antisymmetric. Now, this means that we can contract with the epsilon tensor on three indices, and we won't lose any information. The expression will still equal zero, but now we only have one free index left floating around from the epsilon tensor. Thus, we only have four equations.

17 Problems

These problems are designed to get you interacting with the four-vector notation, and deriving some relatively simple properties of the electromagnetic field.

Problem 1: Electromagnetism as a Requirement of Special Relativity
In your own words, explain qualitatively why magnetism arises from the Coulomb Force Law and Special Relativity. Do not include any equations in your answer.

Problem 2: Time Dilation and Length Contraction
Consider me in my rest frame. I measure the time difference between an event at $(0, 0, 0, 0)$ and $(t, 0, 0, 0)$. Now, boost these events to a rest frame moving at speed v with respect to me. Write the time difference this other observer sees between these two events as a relationship between the observed and proper time.

Next, I see an object moving towards me at speed v. To measure it's length, I record simultaneous events at $(0, 0, 0, 0)$ and $(0, L, 0, 0)$. Calculate where these events occur in the object's rest frame. The object can measure the location of it's two end points at different times, because it's not moving in it's rest frame. Thus, calculate the relationship between the *proper length* of the object and the length that I measure (proper length is the equivalent of proper time – it's the length of an object in the rest frame of that object).

Problem 3: Relativistic Doppler Shift
Consider a photon, of energy $E = hf$, traveling in the x direction. Construct the four-momentum. What happens if we boost in the positive x direction? Construct the Lorentz transformation, and calculate the four-momentum in this frame. How does the frequency change as a function of velocity? Next, boost your original four-momentum in the y direction, and perform the same calculation. You should find that you have a Doppler shift, but your classical intuition shouldn't have expected one (it doesn't occur in sound waves, for example). This transverse relativistic Doppler effect is important in astronomy.

Problem 4: Relativistic Addition of Velocity
Write down a velocity four-vector in the x direction. Now, boost this vector in the x direction by a different velocity. Calculate the resultant three-velocity from doing so. Now, instead of boosting in the x direction, boost in the y direction, and calculate the resultant three-velocity from this process. (Hint: Start with a zero velocity vector, and boost it to get the first velocity vector, using rapidity. When you perform the second boost, you can use hyperbolic trigonometric identities to perform this calculation straightforwardly.)

Problem 5: Longitudinal Electric Field
Starting with an electric field in the x direction, boost in the x direction. What happens to the electric field? What happens to the magnetic field? Explain what you have just derived.

Problem 6: Infinite Charged Plane
Consider an infinite charged plane in the x-y plane with uniform surface charge density. Write down the electric field for this system, and then Lorentz boost it in the x direction. Describe the electric field, and explain if it is what you would expect, based on your original electric field and length contraction. What does the magnetic field look like? (Bonus question for later: Does this agree with Ampere's Law?)

Problem 7: Lorentz Invariants
Calculate $F_{\mu\nu}F^{\mu\nu}$ in terms of the electric and magnetic fields. What is special about this quantity?

Problem 8: Gauge Transformations
Recall that the electric potential has some freedom in where you choose the zero. When you move to a four-potential, the amount of freedom you have available increases. Show that the electric and magnetic fields are unchanged if you add a divergence to the electromagnetic potential, $A^\mu \to A^\mu + \partial^\mu f(x^\nu)$, where f is any differentiable function of spacetime. Calculate how this transformation (known as a *gauge transformation*) affects the electric potential and magnetic vector potential.

Problem 9: Continuity Equation
Consider the equation $\partial_\mu F^{\nu\mu} = 4\pi J^\nu$. What equation do you get if you differentiate noth sides and contract with the free index? Write the resulting equation in terms of the three-dimensional charge density and current. This equation is known as the *continuity equation*, and tells you that charge must be conserved. As a bonus challenge, see if you can figure out how to derive conservation of charge from the equation.